Europäisches Patentamt | **European Patent Office** | **Office européen des brevets**

Bescheinigung Certificate Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

02076345.4

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk

Anmeldung Nr:
Application no.: 02076345.4
Demande no:

Anmeldetag:
Date of filing: 05.04.02
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

Koninklijke Philips Electronics N.V.
Groenewoudseweg 1
5621 BA Eindhoven
PAYS-BAS

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

Signal processing

In Anspruch genommene Prioriät(en) / Priority(ies) claimed /Priorité(s) revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

H04S/

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of filing/Etats contractants désignées lors du dépôt:

AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR

Signal processing

FIELD OF THE INVENTION

The present invention relates to processing of information signals and, more particularly to processing of audio signals.

BACKGROUND OF THE INVENTION

The introduction of new systems like DVB and DVD has brought digital multi-channel sound within reach of a large group of users. The majority of users will however stay for a long time with stereo sound reproduction.

One solution to both serve consumers with 2-channel equipment and multi-channel equipment is so-called simulcast. In this case, two separate information signals are transmitted in parallel, one containing a representation of the multi-channel sound and one containing a representation of the 2-channel sound. To achieve an economical use of transmission or storage capacity, audio bit rate reduction will be used in most applications. The transmitted or stored information signal will then be in the form of a coded bit stream, which requires a decoder to retrieve the audio signal to be reproduced. Nevertheless, it is obvious that simulcast is an expensive solution in terms of required transmission or storage capacity. This makes this solution unacceptable in most practical situations.

Another solution is to transmit only the multi-channel information signal, directly serving the consumers with multi-channel sound reproduction equipment. The 2-channel users then need a decoder that consists of a multi-channel decoder, followed by a downmix module that creates a downmix from multi-channel to 2-channel. Such a 2-channel decoder is thus more complex than a regular multi-channel decoder. In this case, the 2-channel users (the majority) have to pay for the multi-channel capability of others.

It is undesirable that those users are burdened by the multi-channel audio capabilities of a system, in the form of higher costs or higher power consumption. It is also undesirable to waste bandwidth of the system by simulcast (storage and transmission of both a 2-channel (stereo) and a multi-channel stream).

An encoding system that allows a single coded multi-channel audio stream to be decoded by both a true stereo decoder and a multi-channel decoder is the MPEG-2 audio

backwards compatible multi-channel coder (MPEG-2 BC). In all other coding systems, the stereo decoder is basically a (an expensive) multi-channel decoder followed by a down-mix to stereo.

The MPEG-2 BC coder achieves this by performing at the encoder side a down-mix from e. g. 5 channel sound to stereo, coding this as a pure stereo stream, and encoding as an extension three properly chosen signals out of the five input signals. The stereo decoder only decodes the pure stereo stream. A multi-channel decoder also decodes the extra information, and uses an inverse matrix to retrieve the original 5 channels from the down-mix and the additional three channels.

US-B1-6 275 589 describes MPEG-2 having backwards compatibility with MPEG-1, whereby the signals of multi-channel sound channels are matrixed. Stereo signals calculated in a process are then transmitted as an MPEG-1-compatible stereo signal and remaining audio signals are transmitted as supplementary data. This method is known as "compatibility matrixing".

In "Compatibility Matrixing of Multi-Channel Bit Rate Reduced Audio Signals" by ten Kate, preprint 3792, 96th AES Convention, 1994, February 26 – March 01, Amsterdam, it is recognised that the MPEG-2 BC system is not working in an optimal way in case one of the signals in the multi-channel configuration is down-mixed to both the left and right channel of the stereo downmix. This is specifically the case for the Centre channel or for a monophonic Surround channel. The first situation is commonly referred to as the "Dominant Centre" situation.

SUMMARY OF THE INVENTION

It in an object of the invention to provide improved encoding of multiple input signals employing compatibility matrixing.

According to an aspect of the invention, this object can be realised by orthogonalising a composition of signals prior to encoding.

It is another object of the invention to provide dominant centre and/or surround signal processing of multi-channel audio signals, in particular when down-mixing multiple channels to stereo.

According to another aspect of the invention, there is provided sum/difference signal coding of a compatible signal in case of a dominant centre signal situation or a dominant surround signal situation.

According to another aspect of the invention, there is provided a 2-channel decoder that is capable of dominant centre and/or surround processing.

According to a preferred embodiment of the invention, there is provided a method for encoding N input signals, with N>2, said method comprising the steps of:

- generating from the N input signals a composition of M signals, with N>M≥2,
- encoding the composition of M signals into coded data,
- encoding a selection of N-M out of the N input signals into coded data,

wherein the composition of M signals is orthogonalised prior to encoding.

Preferably, M=2.

Preferably, orthogonalisation is done in the frequency domain.

Preferably, orthogonalisation is done by switching between independent coding and sum/difference coding, preferably to sum/difference coding in case of a dominant centre situation.

Preferably, switching between independent coding and sum/difference coding can be selected per frequency band.

According to another embodiment of the invention, there is provided a method for decoding two coded bit streams, the method comprising the steps of

- decoding a first coded bit stream into a composition of M signals,
- decoding a second coded bit stream into a set of N-M signals,
- generating a set of N output signals from the composition of M signals and the set of N-M signals,

wherein the composition of M signals is de-orthogonalised prior to the generation of N output signals.

According to another embodiment of the invention, there is provided an apparatus for encoding N input signals, with N>2, said apparatus comprising means for:

- generating from the N input signals a composition of M signals, with N>M≥2,
- encoding the composition of M signals into coded data,
- encoding a selection of N-M out of the N input signals into coded data,
- orthogonalising the composition of M signals prior to encoding.

According to another embodiment of the invention, there is provided an apparatus for decoding two coded bit streams, the apparatus comprising means for

- decoding a first coded bit stream into a composition of M signals,
- decoding a second coded bit stream into a set of N-M signals,

- generating a set of N output signals from the composition of M signals and the set of N-M signals,
- -de-orthogonalising the composition of M signals prior to the generation of N output signals.

According to another embodiment of the invention, there is provided a signal format for use in transmitting coded data of M signals generated from N input signals, with N>M≥2, comprising:

- a composition of M signals,
- a composition of N-M signals, wherein said composition of M signals is orthogonalised.

These and other aspects and embodiments of the invention will be apparent from the preferred embodiments(s) described hereinafter.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will be more clearly understood from the following description of the preferred embodiments of the invention read in conjunction with the attached drawings, in which:

Fig. 1 illustrates a block diagram of a system in which the present invention is implemented.

Fig. 2 illustrates a signal going out from an encoder after down-mixing.

Fig. 3 illustrates a flow diagram for a method according to a preferred embodiment of the invention.

DETAILED DESCRIPTION OF THE INVENTION

In Fig. 1 it is illustrated an overall block diagram of a system 10 in which the present invention is implemented. The system 10 comprises a matrix 1, an encoder 2 including a stereo encoder 2a and a surround extension encoder 2b, a down-mixing unit 3, a stereo decoder 4, including a stereo decoder 4a and a surround extension decoder 4b, an inverse matrix 5 and means 15 for switching the coding carried out in the encoder 2 between at least two coding modes. The system 10 illustrated in Fig. 1 shows a multi-channel encoder/multi-channel decoder system having down-mix in the encoder.

The encoder 2 can be provided with means for:

- generating from the N input signals a composition of M signals, with N>M≥2,
- encoding the composition of M signals into coded data,
- encoding a selection of N-M out of the N input signals into coded data,

- orthogonalising the composition of M signals prior to encoding.

N input channels, here multi-channel audio signals, a left channel L, a right channel R, a centre signal C, a left surround signal LS, and a right surround signal RS are first transmitted to the matrix 1 and further to the encoder 2 comprising the stereo encoder 2a and a surround encoder 2b. The encoding performed is a so-called "perceptual audio encoding", whereby each of a succession of time domain blocks of an audio signal is coded in the frequency domain. Specifically, the frequency domain representation of each block is divided into coder bands, each of which is individually coded, based on psycho-acoustic criteria, so that the audio signal is compressed efficiently. When the audio signal comprises two or more input channels, typically a left L and a right R channel (stereo), the coding is carried out on a corresponding number of "matrixed" channels L0 and R0. However, in this case five incoming signals are shown. Other types of coding schemes can also be possible, but are not further described in this example.

The encoding can also be performed by software-only solutions.

The encoded signal is down-mixed (or blended) in the down-mixing unit 3 and transmitted as a signal Qout to the decoder 4 as a composition of M signals in a first bit-stream and a selection of N-M signals in a second bit-stream (illustrated as two arrows going into the decoder 4). The signal Qout is illustrated in Fig. 2, which illustrates the two bit streams "onto" each other. Each bit-stream comprises a header 7 and data fields 8.

The decoder 4 can comprise means for:

- decoding a first coded bit stream into a composition of M signals,
- decoding a second coded bit stream into a set of N-M signals,
- generating a set of N output signals from the composition of M signals and the set of N-M signals,
- -de-orthogonalising the composition of M signals prior to the generation of N output signals.

Fig. 3 illustrates a flow diagram of a method according to a preferred embodiment for encoding N input signals, with N>2. In a first step 101, N input signals are transformed to a frequency domain prior to encoding. In a second step 102, a sum/difference mode is selected in case of a dominant centre situation (indicated Yes) or an independent mode in case of no such situation (indicated No). This step is also referred to as the orthogonalisation step. In a third step 103, a composition of M signals is coded into a bit stream of data, typically a first bit-stream and a selection of N-M out of the N input signals is coded into another bit-stream of data, typically a second bit-stream of data.

For instance, the orthogonalisation can be performed per frequency band.

According to another embodiment of the invention, there is provided a method for decoding two coded bit streams, the method comprising the steps of

- decoding a first coded bit stream into a composition of M signals,
- decoding a second coded bit stream into a set of N-M signals,
- generating a set of N output signals from the composition of M signals and the set of N-M signals

wherein the composition of M signals is de-orthogonalised prior to the generation of N output signals. This embodiment can also be illustrated by the flow diagram illustrated in Fig. 3 and will therefore not be further disclosed herein, since "encoding" can be replaced by "decoding".

Examples of matrix equations will be described below to explain the invention better. The matrix equations 1- 21 describe a situation where the present invention is not necessary. These equations are shown to describe the encoding and decoding before describing the equations of a preferred embodiment of the invention for a better understanding of the invention.

Example matrix equations are the following (gain factors are omitted for clarity):

At the encoder side:

$$Lo = L + C + LS \qquad (1)$$

$$Ro = R + C + RS \qquad (2)$$

$$T3 = C \qquad (3)$$

$$T4 = LS \qquad (4)$$

$$T5 = RS \qquad (5)$$

where the transmission channels are: L0 , R0, T3, T4 and T5.

At the decoder side:

$$C' = T3' \qquad (6)$$

$$LS' = T4' \qquad (7)$$

$$RS' = T5' \qquad (8)$$

$$L' = Lo' - C' - LS' = Lo' - T3' - T4' \qquad (9)$$

$$R' = Ro' - C' - RS' = Ro' - T3' - T5' \qquad (10)$$

where the sign ′ denotes a decoded signal.

Although the matrix inversion at the decoder side is exact, the equations above do not yield exactly the original input signals, because the transmission channels L0, R0, T3, T4 and T5 are altered by the encoding. The coding of T3, T4 and T5 is directly controlled by the perceptual encoder and by the quality of C′, LS′ and RS′, which will not give rise to problems.

In the example presented above, due to the matrixing, the coding noise in L0, T3 and T4 will appear in L′, and the coding noise in R0, T3 and T5 will appear in R′. This coding noise could be minimised by choosing appropriate extra channels to be transmitted with L0 and R0. If C, LS and RS are the weakest signals, then the coding noise in L′ and R′ will be dominated by L0′ and R0′, respectively, which is again directly controlled by the perceptual encoder. If another signal combination is the weakest, this signal combination should be chosen to be transmitted as T3, T4 and T5.

However, when the centre signal C is the strongest signal (in the following referred to as the "dominant centre" situation), L0 is almost equal to Ro.

This can be represented by the following formulas:

At the encoder side:

$$L0 = L + C + LS \qquad (11)$$
$$R0 = R + C + RS \qquad (12)$$
$$T3 = L \qquad (13)$$
$$T4 = L \qquad (14)$$
$$T5 = RS \qquad (15)$$

At the decoder side:

$$L' = T3' \qquad (16)$$
$$LS' = T4' \qquad (17)$$
$$RS' = T5' \qquad (18)$$
$$C' = L0' - L' - LS' = Lo' - T3' - T4' \qquad (19)$$
$$R' = R0' - C' - RS' = R0' - C' - T5' \qquad (20)$$
$$= R0' - L0' + T3' + T4' - T5' \qquad (21)$$

where R' is small, $R0' - L0'$ are both large and $T3' + T4' - T5'$ are all small.

It can be shown that one of the small signals always needs to be retrieved by subtracting two large almost equal signals to obtain a small signal. It is clear that a relatively small error in L0 or R0 will lead to a relatively large and clearly audible error in the resulting signal. This quality could be maintained; but only by coding at least one of the compatible signals T3, T4 or T5 at a much higher bit-rate than is necessary for good sound quality of that signal on itself. As described above, another way could be to code additional transmission channels, in this case for instance four, but this is typically a waste of bandwidth as well.

Therefore, according to an aspect of the invention, there is provided an encoder for sum/difference coding of the compatible signal in case of a dominant centre situation. In this way, the centre signal C falls out of one of the equations for the compatible signal, and that equation can be used to calculate a fourth small signal. Of course, for a non-dominant situation, everything can remain the same. For a dominant situation, a matrixing of the compatible signal is added:

At the encoder side:

$$L0 = L + C + LS \qquad (22)$$
$$R0 = R + C + RS \qquad (23)$$
$$T3 = L \qquad (24)$$
$$T4 = LS \qquad (25)$$
$$T5 = RS \qquad (26)$$

$$Ch0 = L0 + R0 = L + R + 2C + LS + RS \qquad (27)$$
$$Ch1 = L0 - R0 = L - R + LS - RS \qquad (28)$$

At the decoder side:

$$L' = T3' \qquad (29)$$
$$LS' = T4' \qquad (30)$$
$$RS' = T5' \qquad (31)$$
$$R' = L' + LS' - RS' - Ch1' = T3' + T4' - T5' - Ch1' \qquad (32)$$
$$2C' = Ch0' - L' - R' - LS' - RS' = Ch0' + Ch1' - 2T3' - 2T4' \qquad (33)$$

Now R′ can be obtained from small signals only, C′ from one strong signal (Ch0′) plus a number of small signals. The situation wherein strong signals are subtracted from each other to obtain a small signal is avoided in this way. In the compatible stereo decoder 4, the following matrix has to be performed:

$$L0 = (Ch0 + Ch1)/2 \quad (34)$$
$$R0 = (Ch0 - Ch1)/2 \quad (35)$$

If the surround extension decoder 4a is omitted, there is provided a multi-channel encoder/stereo decoder instead. In this case, two transmission channels providing a stereo stream carry matrixed surround. The matrix 5 in the decoder 4 is configured so that the two channels do not contain a normal stereo down-mix but a matrixed surround signal. Three signals are then transmitted in an additional signal. Using a matching inverse matrix in the decoder 4 makes it possible to retrieve the original 5 channels from the stereo stream and the additional three channels. In addition, the output of the stereo decoder 4a can be directly transmitted to an input of a matrixed surround decoder.

Another situation where the invention finds application is when one or both of the surround signals LS and/or RS is/are the strongest signal(s). This is referred to as a so-called “dominant surround solution”. In this situation, L0 is in amplitude almost equal to R0 but in anti-phase. Selecting the left channel L, the right channel R and the centre signal C for transmission in the T3, T4 and T5 makes it impossible to retrieve LS and RS with an inverse matrix. It can be shown that always one of the small signals needs to be retrieved by adding L0′ and R0′. The weakest of LS and RS should be selected as the third additional signal. This is illustrated in an example below:

$$L0 = L + C - LS - RS \quad (36)$$
$$R0 = R + C + LS + RS \quad (37)$$
$$T3 = C \quad (38)$$
$$T4 = L \quad (39)$$
$$T5 = RS \quad (40)$$

At the decoder side:

$$C' = T3' \quad (41)$$

$$LS' = T4' \quad (42)$$
$$RS' = T5' \quad (43)$$
$$LS' = L' - C' - L0' - RS' = T4' + T3' - L0' - T5 \quad (44)$$
$$R' = R0' - C' - LS' - RS' = R0' - T3' - LS' - T5' \quad (45)$$

Due to the fact that L0′ and R0′ are in anti-phase this means that adding two large almost equal signals to obtain a small signal. It is clear that a relatively small error in L0′ or R0′ will lead to a relatively large and clearly audible error in the resulting signal. The quality can still be maintained, but only by coding at least one of the compatible signals with a much higher bit-rate than necessary for good sound quality of that signal on itself. Also in this case could another way be to code additional transmission channels at the cost of waste of bandwidth.

According to another preferred embodiment of the invention, a matrixing of the compatible signal is added according to the following equations:

At the encoder side:

$$L0 = L + C - LS - LS \quad (46)$$
$$R0 = R + C + LS + RS \quad (47)$$
$$T3 = C \quad (48)$$
$$T4 = L \quad (49)$$
$$T5 = RS \quad (50)$$

$$Ch0 = L0 + R0 = L + R + 2C \quad (51)$$
$$Ch1 = L0 - R0 = L - R + 2LS - 2RS \quad (52)$$

At the decoder side:

$$C' = T3' \quad (53)$$
$$L' = T4' \quad (54)$$
$$RS' = T5' \quad (55)$$
$$R' = Ch0' - L' - 2C' = Ch0' - T4' - 2T3' \quad (56)$$
$$2LS' = L' - R' - 2RS' - Ch1' = T4' - R' - 2T5' - Ch1' \quad (57)$$

Now R′ is obtained from only small signals, LS′ from one strong signal (Ch1′) plus a number of small signals. The situation that strong signals are subtracted from each other to obtain a small signal is avoided in this way. In the compatible stereo decoder, the following matrix has to be performed:

$$Lo = (Ch0 + Ch1)/2 \qquad (58)$$

$$Ro = (Ch0 - Ch1)/2 \qquad (59)$$

The invention finds application for instance in CD-2, and other multi-channel music distribution.

The coded data can be stored and subsequently read, decoded and presented to a listener of a record carrier.

It should be noted that the above-mentioned embodiments illustrate rather than limit the invention, and that those skilled in the art will be able to design many alternative embodiments without departing from the scope of the appended claims. In the claims, any reference signs placed between parentheses shall not be construed as limiting the claim. The word 'comprising' does not exclude the presence of other elements or steps than those listed in a claim. The invention can be implemented by means of hardware comprising several distinct elements, and by means of a suitably programmed computer. In a device claim enumerating several means, several of these means can be embodied by one and the same item of hardware. The mere fact that certain measures are recited in mutually different dependent claims does not indicate that a combination of these measures cannot be used to advantage.

CLAIMS:

1. A method for encoding N input signals, with N>2, said method comprising the steps of:

- generating from the N input signals a composition of M signals, with N>M≥2,
- encoding the composition of M signals into coded data,
- encoding a selection of N-M out of the N input signals into coded data,

wherein the composition of M signals is orthogonalised prior to encoding.

2. A method according to claim 1, wherein the composition of M signals is coded into a first bit-stream, and the selection of N-M signals is coded into a second bit-stream.

3. A method according to claim 1 or 2, wherein M=2.

4. A method according to any one of the claims 1-3, further comprising the step of:

- selecting of a sum/difference mode versus an independent coding mode.

5. A method according to any one of the claims 1-4, wherein the N input signals is transformed to a frequency domain prior to encoding.

6. A method according to any one of the claims 1-5, wherein the orthogonalisation is performed per frequency band.

7. A method for decoding two coded bit streams, the method comprising the steps of

- decoding a first coded bit stream into a composition of M signals,
- decoding a second coded bit stream into a set of N-M signals,
- generating a set of N output signals from the composition of M signals and the set of N-M signals,

wherein the composition of M signals is de-orthogonalised prior to the generation of N output signals.

8. Apparatus for encoding N input signals, with N>2, said apparatus comprising means for:

- generating from the N input signals a composition of M signals, with N>M≥2,
- encoding the composition of M signals into coded data,
- encoding a selection of N-M out of the N input signals into coded data,
- orthogonalising the composition of M signals prior to encoding.

9. Apparatus for decoding two coded bit streams, the apparatus comprising means for

- decoding a first coded bit stream into a composition of M signals,
- decoding a second coded bit stream into a set of N-M signals,
- generating a set of N output signals from the composition of M signals and the set of N-M signals,
- de-orthogonalising the composition of M signals prior to the generation of N output signals.

10. A signal format for use in transmitting coded data of M signals generated from N input signals, with N>M≥2, comprising:

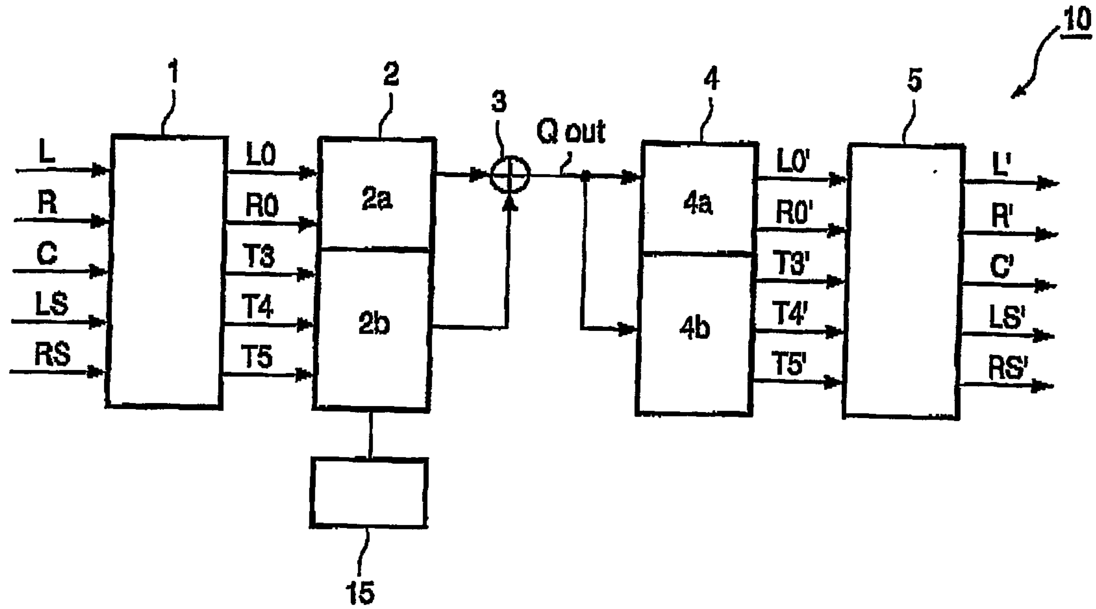
- a composition of M signals,
- a composition of N-M signals, wherein said composition of M signals is orthogonalised.

11. A record carrier on which a signal format as claimed in claim 10 has been stored.

ABSTRACT:

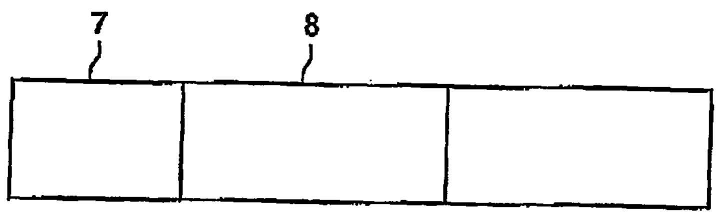
Sum/difference coding of a compatible signal, typically in case of a dominant centre signal or dominant surround situation of a multi-channel audio stream to be decoded by both a stereo decoder and by a multi-channel decoder, to provide improved encoding of multiple input signals employing compatibility matrixing.

Fig. 1



10
1
2
3
4
5
Q out
L
R
C
LS
RS
L0
R0
T3
T4
T5
2a
2b
4a
4b
L0'
R0'
T3'
T4'
T5'
L'
R'
C'
LS'
RS'
15

FIG. 1



7
8

FIG. 2

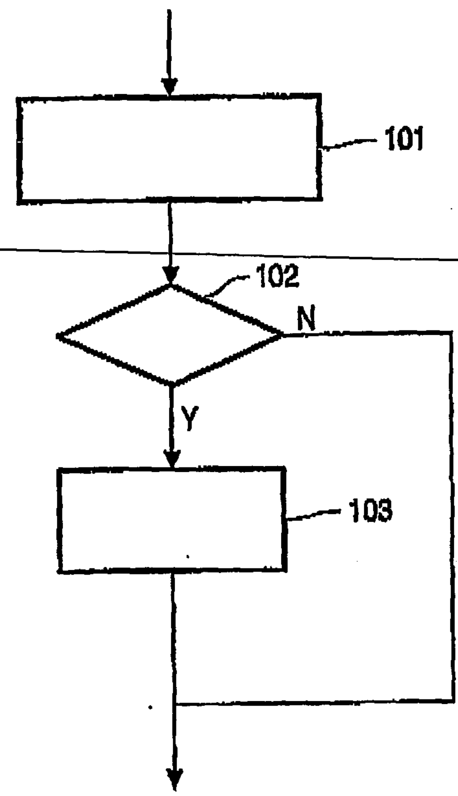


FIG. 3